Effects of County-Level Diversity on Educational, Economic, and Health-Related Outcomes

Ken Groszman, Jack Petersen, Bing Xu– Section 01, Group 6
Final Project for STAT 405– R for Data Science, Dr. Roberto Bertolusso
Department of Statistics, Rice University

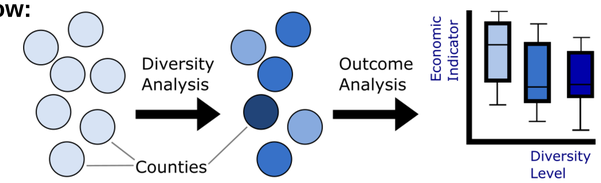
RICE

Introduction

Motivation:

- In today's society, it is common to hear the government, businesses, and members of our community touting the benefits of diversity [1]
- Indeed, diversity provides micro-scale benefits in certain organizations (new perspectives, increased productivity, expanded worldview, etc.) [2]
- We seek to **examine macro-level effects of county-level diversity** on a variety of downstream indicators of success: education, economics, and health

Workflow:



Calculation of Diversity Index

Description of Methods:

- Each county *j* is assigned a diversity index value d_j based on the race breakdown of its population
- The diversity index is calculated according to the following formula [3], where x_{ij} is the proportion of race *i* in county *j*:

$$d_j = 1 - \sum_{i \in R} x_{ij}^2$$

- A county with the highest possible d_j has its population spread exactly evenly across the 8 races, whereas the lowest possible score has its entire population as one of the groups (i.e. 100% white, 0% everything else)

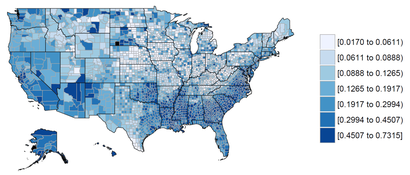
Data:

- Data is pulled from IPUMS (Integrated Public Use Microdata Series), which is census-level data for selected counties in the USA

Variableo	Description	Source
Geographic Location	County location of record, given by state and county FIPS (Federal Info. Processing Standard) code	[4]
Race	Numerical code corresponding to race of individual: white, black, American Indian, Chinese, Japanese, Pacific Islander, other, 2+ races	[4]

Results:

US County Diversity Choropleth



- Southeast and southwest USA, Alaska, quite diverse
- North, Midwest, and Central USA among less diverse regions

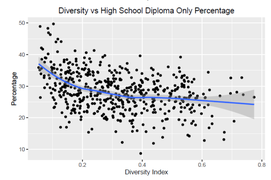
Educational Outcomes

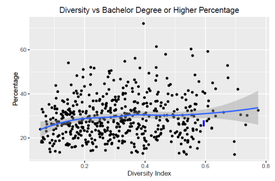
Here, we look for correlations between county diversity and educational outcomes

Data:

Variable	Description	Source
Educational Attainment	Code corresponding to highest level of education achieved: no high school, high school diploma, Associate's degree, Bachelor's degree,	[5]

Results and Interpretation:





More diverse counties have a higher percentage of people reaching the highest education level, and that less diverse counties have the most number of people with low educational attainment (high school only)

Economic Outcomes

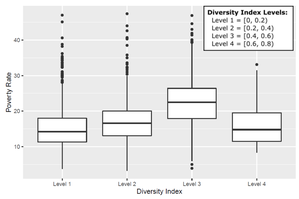
Here, we look for correlations between county diversity and economic outcomes

Data:

Variable	Description	Source
Poverty Status	Code indicating whether individual falls below the poverty line	[6]
Income	Integer indicating the individual's yearly income	[6]

Results and Interpretation:

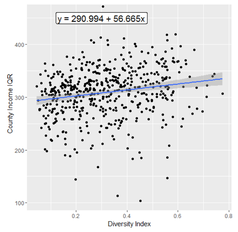
Poverty Rates for Counties in Different Diversity Index Ranges:



- In intermediate diversity levels, possibly minority groups marginalized leading to higher poverty rates
- At very low diversity and at very high diversity, this phenomenon not observed

Effects of Diversity on County-Level Income Gap:

- Calculated income gap as the IQR of household incomes for a given county
- Higher diversity index is correlated with a higher income gap
- Correlation weak, since other factors important (location, main industries, etc.)
- Trend could be explained by racial divide between highest and lowest-paying jobs in diverse counties



Health-Related Outcomes

Here we look for correlations between county diversity and health-related outcomes

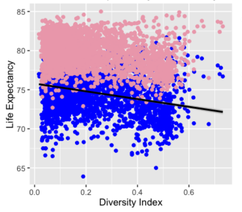
Data:

Variable	Description	Source
Life expectancy	Database of life expectancy for different counties in the US	[7]
Obesity	Rate of obesity in different US counties	[7]

Results and Interpretation:

*****In plots, blue dots represent males and pink dots represent females*

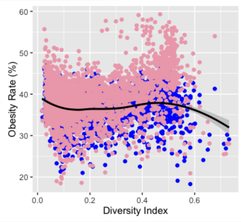
Variations in County-Level Life Expectancy as Result of Diversity



- Life expectancy decreases with increased diversity
- Potential causes:
  - More diverse counties urban → higher risk of disease
  - Certain racial groups have lower life expectancies, driving down the number for those counties

Variation County-level Obesity Rates with Diversity

- Obesity rates tend to be highest in counties with medium-high diversity
  - Possibly correlated with higher poverty levels → people eating more junk food
  - More data necessary to discern whether effect is caused by diversity or just correlated with it



Conclusions & Future Research

Conclusions:

- County-level diversity is indeed correlated with educational, economic, and health-related outcomes
- The analysis is limited, since many other external factors also contribute to these outcomes (location, major industries, size of county, urban vs. rural, etc.)

Future Research Directions:

- Currently, our analysis is limited by access to data
  - Would have been good to control for race (i.e. see how a certain race does in counties with different diversity indices)
  - Could also look at data in reverse order (of people with high income, are they generally from counties with higher or lower diversity)

References

[1] "Are There Really Any Tangible Benefits to Diversity and Inclusion? Isn't This Just Corporate Jargon?" Government Agency. *US Office of Personel Management*. Accessed November 28, 2016. https://www.opm.gov/faqs/QA.aspx?fid=72bcd219-0b9f-4de8-b366-4817028fbc6e&pid=f2ef3151-b4f2-4f47-a319-acad8175b0b7.
[2] Abreu, Kim. "The Myriad Benefits of Diversity in the Workplace." *Entrepreneur*, December 9, 2014. https://www.entrepreneur.com/article/240550.
[3] "Diversity Index of US Counties | Kaggle." Accessed November 30, 2016. https://www.kaggle.com/mikejohnsonjr/us-counties-diversity-index.
[4] Ruggles, Steven, Katie Genadek, Ronald Goeken, and Matthew Sobek. "Integrated Public Use Microdata Series: Version 6.0 [Machine-Readable Database]." Minneapolis: University of Minnesota, 2015. https://usa.ipums.org/usa/.
[5] "2015 US County Race Data." Government Agency. *US Census Bureau*. Accessed November 30, 2016. https://www.census.gov/popest/data/counties/asrh/2015/files/CC-EST2015-ALLDATA.csv.
[6] "IHME US Data for Download." Government Agency. *Institute for Health Metrics and Evaluation*. Accessed November 30, 2016. http://www.healthdata.org/us-health/data-download.
[7] "USDA-ERS County-Level Datasets." Government Agency. *United States Department of Agriculture Economic Research Service*. Accessed November 30, 2016. https://www.ers.usda.gov/data-products/county-level-data-sets/download-data.aspx.